

QY	403	NTQTEKILAGVGQOR-----EGVILIRASQEOIRLETITDDSSRIIMHIRRGSESSRG	454
		: : : : : : :	
Db	236	NTQREKLEILIEBGCGAKRQGGGKPFKAPEDIRKISQDAISPRH-----RGE-RLA	34.9
QY	455	YNLEFKRLYSNKYGOAYEVKPEDYROLQDMDSVLTANTYQGSMMGPFENTRSTKVVV	514
		: : : : : : : : : : : :	
Db	350	INLTSQSPVYSNONGRFFEPACPEDFSQFQNDVAVSAFKLNQGIIFVPHYSKATFVVF	409
QY	515	ASGEADYEMACPHILSGRRIGRGCGK--RHHEEDV-----HYQVYARLSKRALYVLA	566
		: : : : : : : : : : :	
Db	410	TDGIGYVAMACPHILSKRQSGQSGQSRQDRDEDEESEETGEFQVYAPLSPGVFAPA	469
QY	567	GHPVYVSSGNGENLLTPAFGINAQNHNENFLAGREKRVNLIQIQEQAEMELFAAPRKEVEE	626
		: : : : : : : : : : : :	
Db	470	GHAYTFASKQQLPILNAAYAGLNAQNNORIFLAC-KVLVQRMQSEAKELSGVPSKLYDN	528
QY	627	SFNSODOSIFFFGPRHQOQSPBSTKQQQQLPYLTLDP	663
		: : : : : : : : : : : : :	
Db	529	IFNNDESYFMFSQSORQ---RDRKGNPLASTLDF	562

```

RESULT 2
US-07-955-905A-22
: Sequence 22, Application US/07955905A
: Patent No. 5770433
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
: TITLE OF INVENTION: PRECURSOR
: NUMBER OF SEQUENCES: 28
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/955,905A
: FILING DATE: 21-JAN-1993
: CLASSIFICATION: 435
: INFORMATION FOR SEQ ID NO: 22:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 566 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Theobroma cacao
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..566
: OTHER INFORMATION: /note="67 kD Precursor Protein"
: US-07-955-905A-22

```

Query Match	31.8%;	Score 1127;	DB 1;	Length 566;
Best Local Similarity	40.4%;	Pred. No. 7.6e-95;		
Matches 233;	Conservative 109;	Mismatches 177;	Indels 58;	Gaps 13

QY	109	ICEEEENR---QRDQOQOEOQKQCRRETERPRIMOQSOORCERNYKEXKROOKY	165
Db	22	LCSEVSAAGRRQYDRPRQOYEOQRCSEBATEEEOEOECORCER-----EX	70
QY	166	EEQOREDEBEKYEEMKFEEDNKRDOQOREYEDCRRCEOE--PROOQOCLRCREOORH	223
Db	71	KEQDRQOEEL-----QOYQOOCQRCOEQOQOEQOQOQCRKMEQYKEQ	116
QY	224	GRGDDMNPOFGSGRYEEGEEOQSDNPYFDE-RSLSTRPRTTEGHIIVLENFYGRSKL	282
Db	117	ER-GEHENYHNHKKNRSEBEEGQRRNNPYEPKRRSQYFGRDEBGNFKILOFAENSPR	175
QY	283	LRLAKNRYLVLENRPAYLPTNLDADAILLVITGCGALKMTHNDNRSEYULECGDVR	342
Db	176	LKGINDRYLAFMEANPMTFLPHNCDSEALVYFNGSGTTEVTHKEKESYNQORGVVS	235

```

Oy 343 IPGTTPELLINRONNERLHIKFIQTJSTTSPQXQKEFPAGQONPEPLTFSKELLIAL 402
Db 236 VPAGSTVYVVSQDMOERLTLVIALPVNSPCKYTLFFPAGNNKRESTVYGASTVELTVF 2959
Oy 403 NTQTEKLGVEGQOR-----EGVITIRASQEOIRELTRDSESRIHMIIRGESSRGP 454
Db 296 NTQREKTEELLEEOGROGRQOOGQGMERKAPEDQIRAIISOATSPRIH---RGE--RLA 349
Oy 455 YNLFKRPRLYENKTKGOAEVKEPELYROLOMDLSVFLAANTQSGMMPFENFSTKVVVV 514
Db 350 INLTSQSIVYINONGRFFELACPEDEPFSQFQNDVAVSFKLNGALFVPHYNSKATFVVFV 4099
Oy 515 ASGEADYEMACPHLTSGRHGRGSGK--RHHEEDV-----HYEQYRRLSKREAVIYLA 566
Db 410 TIDGCGYAOACPHLSROSQSGOSQSORORREOESESESEETGFEFOOVYAPLSPDVFAPA 469
Oy 567 GHYPVYFVSSGGENLLFLFPGFNAQNNHNFLLAGREERVYLOOIFEOAKELFLFAPRKVEE 626
Db 470 GHAYTFEFSKQOPLNVAVFGLNQNNQRIFLACG-KULVYQMDSEAKELSTFGVPSKLYDN 528
Oy 627 SFNSODOSIFFPGRPHOQOSPRSTKQOOQLVSLTDF 663
Db 529 IENNDESYFMSFSQOORQOR---RDERKGNPLASTLDF 562

```

```

RESULT      3
US-07-955-905A-23
: Sequence 23, Application US/07955905A
: Patent No. 5770433
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
: TITLE OF INVENTION: PRECURSOR
: NUMBER OF SEQUENCES: 28
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/955,905A
: FILING DATE: 21-JAN-1993
: CLASSIFICATION: 435
: INFORMATION FOR SEQ. ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 587 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Gossypium hirsutum
: FEATURE:
: NAME/KEY: Protein
: LOCATION: 1..587
: OTHER INFORMATION: /note="Vicillin from G. hirsutum"
US-07-955-905A-23

```

OTHER INFORMATION: /note= "Vicillin from G. hirsutum"
7-955-905A-23

	Query Match	30.4%	Score 1075;	DB 1:	Length 587;
	Best Local Similarity	38.6%	Pred. No. 4.6e-90;		
	Matches 231; Conservative 113; Mismatches 181; Indels 74; Gaps 14;				
QY	76 EDPOECOCOCORCCKOQESGPRQOYOCCORRC-EICEEEEFNRORDRDOOYECSOKHCQ	134			
	::::: ::::: : ::::: : ::::: : ::::: : ::::: :				
Db	35 DDPKRYEDCSRRCSEMDTRGOKEQQQCESCKSYGEKKQDGNHREDRPQRYVEEQDEC-	93			
QY	135 RRETFPRMOTQOQCSCERRYEKKKKQQRKYEDQREDEEKEYEEMKEEDNKRDPQARY	194			
	: :				
Dd	94 -RQGEERGRPQQQNC-----IKRFQDQDQD-----SQRF	123			
QY	195 EDCRRCEQOE--PROQHOCILCRFEQDRONHGSGDMNPQ-----GGSGRYEGEDE	246			
	::::: ::::: : ::::: : ::::: : ::::: : ::::: :				

```

Db 124 QECQCHQHQQQREPERKQQQVAAECRERYOE-----NPMRRREBEADEEETEEGEQE 175
Qy 247 OSDNYTYDEDSLSIRFTTEGCHLSYLENFGSKLLRALKNIRVYLENPNPAVLPYTH 306
Db 176 QSHNEFHRRSFOSFRHEEGNFRVLOREFASRHPILRGINEFLSLLENPMTEVLPYH 235
Qy 307 LDADAILLVIGRCALKMIHNDNESYNLECGDVIIRIPACTFTFLINRONNERLHIAKFL 366
Db 236 CDAEITYLVTKRGTLITFLTHENKESINVPGVVYRPAASTIYLAQNQKELITIAVLH 295
Qy 357 QTISTPGQYKEFFPAGGONPEPYLPSTFESKELLEAALNTQETKLGVCYFG-----QOREG 419
Db 296 RPNVNPQGFEEFPAGSQRPQSYLARSREILPEAFNTRSEQIDELFEGGQSHRRQGGG 355
Qy 420 VTIASQDQIELTRDDSESHHHIIRGGESSKCPYVLFMKRPLYSKTYGOATEVVKPEDY 479
Db 356 MFRKASQDQIHALSOEATISPR---EKSGE--RFAFNLILRYTPYSNQNRFYEACPREK 409
Qy 480 ROLDMDISYFIANTVGTGSGMMGPPEFNRTSTKVVVYVAGDEADVEACPHLSGRHGGGKG 539
Db 410 RQLSLINVYSALDQNGSITVPHYNSKATIFPVLYNAGNGCYEMVSHLPDROSSFEEBE 469
Qy 540 RHEEEDV-----HYEQVARRLSKREAIYVLGHPVYFVSSGNEMLITPAFG-----IN 588
Db 470 QOQEOEOEERRSQGYRKIRSQLSRGDIFVVPANFPVTFVAAQONQNLBMFGFLYNQIN 529
Qy 569 AONHNENLAGRENVIYDQIEPQAMELAFAPAKREVEESNSODOSITFFPGPRHQOQOS 647
Db 530 PDHNRQIFVACKINHV--RQWDSQAKELAFGVSSRLYDEIFNPNNDPQSEYFVS--RQORAS 586

RESULT 4
US-07-955-905A-24
Sequence 24: Application US/07955905A
Patent No. 5770433
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
TITLE OF INVENTION: PRECURSOR
NUMBER OF SEQUENCES: 28
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/955,905A
FILING DATE: 21-JAN-1993
CLASSIFICATION: 435
INFORMATION FOR SEQ. ID NO: 24:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Glycine max
FEATURE:
NAME/KEY: Protein
LOCATION: 1..605
OTHER INFORMATION: /note= "Viciilin from G. max"
US-07-955-905A-24

```

	Query Match	23.7%	Score 838.5	DB 1:	Length 605;
	Best Local Similarity	30.5%	Pred. No. 2,2+68;		
	Matches	205;	Conservative 144;	Mismatches 223;	Indels 101; Gaps 177;
OY	11 LFLFLSLFLLTYS--LAESFFDROELEEBCRCOCMOLETSCQMRRCYSQCCKREEDID	68			
	: : :				
Db	8 LTLGLVYLASVSSEFGAIYWKEKPKNNKKLQSCNSERDSIRNACIARC-----N	59			
YC	69 WSKYDNODEPOTECOCRCRCRGOESGFRPOOYCRCRKETIEEEENNRDD---PPQQ	125			

[illegible]

RESULT 5
 US-07-955-905A-25
 Sequence 25: Application US/07955905A
 Patent No. 5770433
 GENERAL INFORMATION:
 APPLICANT:
 TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
 TITLE OF INVENTION: PRECURSOR
 NUMBER OF SEQUENCES: 28
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/07/955,905A
 FILING DATE: 21-JAN-1993
 CLASSIFICATION: 435
 INFORMATION FOR SEQ. ID NO.: 25:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 571 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: *Platanus sativum*
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 1..571

OTHER INFORMATION: /note="Conviclin from P. sativum"
US-07-955-905A-25

Query Match 23.2%; Score 821.5; DB 1; Length 571;
Best Local Similarity 31.0%; Pred. No. 7,2e-67;
Matches 211; Conservative 112; Mismatches 205; Indels 153; Gaps 19;

```

QY 11 LFLSLFLSTVSLAESEFDRQYECKRCOCMOLETSQOMRCVSOCDKREEDIDW-69
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 12 LFLGLIILASVCVTA-----NYDEGS---ETRVPGRRRGQEGEKKRNGEMR 60
QY 70 SKTDNEDPOTECQCCORCRQOESCPROOQYCRCKELCEEEERYNRQDPQOQTEOC 129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 PSYEKEEH-----EKKYRYOREKKEQ-----84
QY 130 QKHCQRETEPRIMQTCQCCERREYERKRRKQOQRYEEOQREDEEYEEEMKEDKNDP 189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 -----KEVQPG-----RRWERE-----EDEQVEEEMKQSORREDP 116
QY 190 QOREYEDCRRRCBOEPROOHQCLRCQEOQRHGRGDMNPORGSGRYEEGEESQSD 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 117 EER-----ARLHREERTKDR-----RHQREGEEERSSQSEHR 152
QY 250 NPVEPDERSLSTRFTEEGHISYLENFYGRSKLLRALKNRYLLANPNAYVLP7HLD 309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 153 NPELFKSNKFLTFENENGRIIRLOFDRKSDLEFENLQNYRLVEYRAKPTTILPOHIDA 212
QY 310 DALLVIGRGALKMIHNDRESYNLECGDVIRIPAGTFFYLLINRDNNEHLIAKFLQTI 369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 213 DILLVVLNGKAILTVLSPNDRNSYNLEKGTIKIPAGTTSYLVNODDEDLRVVDVPIV 272
QY 370 SPFGQKEFFPAGQNPPEYLTSTESKEILEALNTQTEKLRGYE-----413
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 273 NRGKFEAF--GLSEKKNQYLRGFSKNILEASINTKYETIEKVLLEEQEKKPOQLDRKR 330
QY 414 ----GOOREGVIITASQOIRLTDSESRHWHIRRGESSRGPNYLFNKRLYSNKYG 470
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 331 TOGGEERD-AIKVSREPOIELEKLAKSSS---KSLSEFEFPNLSIKPEYSKMGFK 385
QY 471 AVEVKE-DYROLQDDMLSVFIANVTQSGMGPFFNTSTKVVVAVSGADVEMACP7HS 529
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 386 LEFTTEPKTPQLODDIDIVSCVEIKNGALMLPHYSRAIVLVLLVEGKNLILL-440
QY 530 GRHGGGCGKRHEEEDV---DVHYEQVRLSKREALIVLAGHPVYVSSGNENLL 582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 441 -----GLKNQOQREDRKERNNEQVREARLSPGDVYIIPAGHPVAISASNLIL-491
QY 583 FARGIAQNNHNEFLAGRENVLOQIEPQAMELAFAPRKVEEESNSQDOSTFFPG-PR 641
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 492 -GFGIAKNNQNRFLSGSDNVISQIENPYKELTFPGSSQEVNRLIKOKSHFASAEPE 550
QY 642 OHQOQSPRSTRKQOQPLVSIID 662
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 551 QKEESOR---KRSPLSSVLD 568

```

RESULT 6

```

US-07-955-905A-26
; Sequence 26, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A

```

```

; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORGANISM: Pisum sativum
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..410
; OTHER INFORMATION: /note="vicillin from P. sativum"
US-07-955-905A-26

```

Query Match 19.7%; Score 699.5; DB 1; Length 410;
Best Local Similarity 36.9%; Pred. No. 6,4e-56;
Matches 147; Conservative 79; Mismatches 143; Indels 29; Gaps 5;

```

QY 249 DNPVYFDERSLSTRFTEEGHISYLENFYGRSKLLRALKNRYLLLEANDPNAYVLP7HLD 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 20 ENPFITKSNRPOTLYENENGRIIRLOFDRKSKIFENLQNYRLLEYKSRPHILFLPQITD 79
QY 309 ADAILVIGRGALKMIHNDRESYNLECGDVIRIPAGTFFYLLINRDNNEHLIAKFLQTI 368
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 80 ADLIVVLSGKATLVKLSNDRNSFNLERQDAIKLPAGSIAYFANDDDEPRVLDLAI 139
QY 369 ISTPGQKEFFPAGQNPPEYLTSTESKEILEALNTQTEKLRGYEFOQR-----417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 140 VNKPGQLQSLFLSGTONKSSLSGFSKNILEAFNPNYEIEKVLLEQOQEPORHRSK 199
QY 418 -----EGVIRASQOIRLTDSESRHWHIRRGESSRGPNYLFNKRLYSNKYG 469
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 200 DRQELINEENVLYKVSROQIEELSKNAKSSS---KSYSSSEGGPNLNSRPNTYSNKG 255
QY 470 QAVEKPEYDRILODDMLSVFIANVTQSGMGPFFNTSTKVVVAVSGADVEMACP7HS 529
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 KFEFTPEKNOQLODDIDIVNSVDIKVGSILLPNYSRAIVTVYTEGKDFELVQGR--313
QY 530 GRHGGGCGKRHEEEDV---HYEQVRLSKREALIVLAGHPVYVSSGNENLLFARGI 587
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 314 NENQKENDKEEEOEETSQVQVLYRAKLSPGDVYIIPAGHPVAISASDLNLI--GLGI 371
QY 588 NAQNNHNEFLAGRENVLOQIEPQAMELAFAPRKVEE 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 372 MAENNERNFLAGEDNVISOVERPYKELAFPGSSHEVD 409

```

RESULT 7

```

US-07-955-905A-27
; Sequence 27, Application US/07955905A
; Patent No. 5770433
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: RECOMBINANT 47 AND 31 KD COCOA PROTEINS AND
; NUMBER OF SEQUENCES: 28
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/955,905A
; FILING DATE: 21-JAN-1993
; CLASSIFICATION: 435
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 421 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

```

MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Phaeolus vulgaris
FEATURE:
NAME/KEY: Protein
LOCATION: 1..421
OTHER INFORMATION: /note= "vicillin from P. vulgaris"
US-07-955-905A-27

Query Match 15.3%; Score 542; DB 1; Length 421;
Best Local Similarity 33.1%; Pred. No. 1.7e-41;
Matches 141; Conservative 71; Mismatches 158; Indels 56; Gaps 13;

QY 244 EEESSDNPYED-ERSSTPTEEGHISVLENYGRSKILRALKNRYLVLEAPAFV 302
||| ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 30 EEESSDNPYED-ERSSTPTEEGHISVLENYGRSKILRALKNRYLVLEAPAFV 302
||| ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 303 LPTHLADAILLVGSGALKMIHNDRESY-----NLECGDVIIPAGTTFYLINRDN 356
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 90 LPOQADAEILLVYRSGSAILLVKPPDRREYFLLTSDNPIFSDHOKIPAGTIFYLVNPD 149
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 357 NEHLIAKELQITSTPQYKEFPAGQNPPIYLSKTEALNTOTKLRGVFGQ- 415
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 150 KEDLRITQLAMPVNP-QIHEFLSTEAQOYLOEFSKHLLEASFNSKEEINRYLFE 208
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 416 --QREGYIIRASQOIRELTDSESRHMHIRGSGSRGPYNLFNKRPLYSKNGAYE 473
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 209 EGQEGYIVIDEQIKELK-----HAKSSRKS-----LSKQDNTIGDFGLTE 255
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 474 VKBEDYROLQMDLSVFIANV--TQSGMGPFFNTRSTKVVVAAGADVEMACPILSGR 531
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 256 -----RTDNLNLVLISSIEHEGALFVPHYKSAIVLVNVEGAHVELGPK----- 303
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 532 HGRGGRGKRHEEDVYEVQARLSKREAIYVLGHPVYVSSGNLLFAFGIYAON 591
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 304 -----GNK-----ETLEYESYRAELSKDQVYIPAAVPAIKATSNVNF--GEGINANN 351
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 592 NHENFLAGERNVLOQI-----EPQAMELAFAPRKVEESFNSQDSIFPPGRHOQO 646
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 352 NNRLLAGKTDVNISSIGRALDQKDVIGLTFSSGDEVMLINKOSS-YFVDAAHHQOE 410
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 647 SPRSTK 652
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 411 QQKGRK 416

APPLICATION NUMBER: US/08/056,200
FILING DATE: 30-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fedrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054, 001A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 94:
SEQUENCE CHARACTERISTICS:
LENGTH: 1898 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-056-200-94

Query Match 6.1%; Score 216.5; DB 1; Length 1898;
Best Local Similarity 20.8%; Pred. No. 1.1e-10;
Matches 154; Conservative 110; Mismatches 289; Indels 189; Gaps 30;

QY 28 ESEPDQYEYEC-----RQOMLETSG-----QMRVCSDCKRF--EE 65
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 185 ERRAEEQLOSCGHETEERPDDEQLRRLELRLRKGRGKQOQR--EQDVRPQEE 242
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 66 DIMSKTD-----NODPOTECQ-----QOCRCROQSGP-----R 97
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 243 EKEMRKRETVLRKEEKLQEEEPQRELOEEERQOLKRLERREOREEEOQOQLR 302
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 98 OOOYCORCKEICEEERERORPOOYBOCOKHCORRETERPHMOTC-----QQ 148
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 303 EQOLRRKQEEERQOEEERQOEEERQOEEERQOEEERQOEEERQOEEERQO 362
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 149 RCERRYEKERKQOKR--YEEQKEDKYEERKEDN--KRPQOREVEDCR--RC 201
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 363 QLRQEEERQOEEERQOEEERQOEEERQOEEERQOEEERQOEEERQOEEERQO 422
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 202 EQOEPQO--HQCQLRCRQO--KQHGSGDMANPQSGRYEGEEEOSD----- 249
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 423 EQOLRRQOOLRRQOOLR--RQOOLRRQOEEERHCKQEEERQOLRRQOEEERQO 481
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 250 -NPYFDERSLSTRFTEEGHIS--VLENFYGRSKILRALKNRYLVLEAPNPAFLP 304
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 482 ETERHDEERKQOLKRDQEEERERKWLKEEERREQOEEERQOOLRRQOEEERQO 536
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 305 THLDADAILLVIGRGALKMIHNDRESYMLECGDVIIPAGTTF-YLINRDNNERLIA 363
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 537 -----RLKQEEERRLQORLRSQOLRRQOEEERLQOLRRQOEEERLQOE 580
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 364 KFLQITSTPGQYK--EFPFAGQNPPIYLSKTEALNTOTKLRGVFGQ- 409
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 581 RRQOLRRQOEEERQOOLRRQOEEERQOOLRRQOEEERQOOLRRQOEEERQO 640
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 410 RGVFGQOREGVIIIRASQOIR--ELTRDSESRHMHIRGSGSRGPYNLFNKRPLYSN 466
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 641 REEPERERHLLKSEQOEEERHQLRRQOEEERQOOLRRQOEEERQOOLRRQOEEER 698
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 467 KYQOAYEVKPEDYROLQMDLSVFIANVNTQSGMGPFFNTRSTKVVVAAGADVEMACP 526
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 699 RRDQ--ELAEEDQOARER-----IKSRIPKQWQOLESEADARQSKV 738
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 527 HLSGRHGGG-----KRHEEDVYEVQARLSKREAIYVLGHPVYVSSGNENL 560
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 739 LLEAPQARAPQOEQEKRRSELOQOEERARHQOE-----EQR 782
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 581 LIFAFGINQNNHENVLAGERNVLOQIEPQAMELAFAPRKVEE-----SPNSQDSI 635
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
Db 783 RDTWQOAEKSR--GRQ-----LSARPLRQOREOLRAEERQOER 827
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :
QY 636 FPPGPRHOQOOSPRSTKQOQPL 657
||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : : ||| ||| : : : : :

Db 274 RRQKRRRMEQERIRQEYD---EKERQRY 300

RESULT 15

US-08-458-860A-2
; Sequence 2, Application US/08458860A
; Patent No. 6100390
; GENERAL INFORMATION:
; APPLICANT: Frank, Glenn R.
; APPLICANT: Tripp, Cynthia A.
; APPLICANT: Griever, Robert B.
; TITLE OF INVENTION: NOVEL PARASITIC HELMINTH
; TITLE OF INVENTION: P220 NUCLEIC ACID MOLECULES
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,860A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 336
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-13-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 303 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-458-860A-2

Query Match 4.0%; Score 143; DB 3; Length 303;
Best Local Similarity 24.6%; Pred. No. 3.6e-05;
Matches 52; Conservative 46; Mismatches 87; Indels 26; Gaps 10;

QY 32 DROVEECKRQ---CMQLETS---GOMRRC--VSOCDKRPEE-----DIDMSKYDNOEDP 78
DB 97 DDEDLDECSQDERFCRYLAKTLGVHLKICDGDIDCGSDSEMCADDEYITISINNEST 156
QY 79 QTECCQCCQRCRCRQESGP--RQOQYCOR--CKEICEEE--EYNRQRPQOQYECCQKHC 133
DB 157 NIRCDDPDQFRC---ENGKCIAGIDRCNRKYDCDDGTDETTCEYFQAIOQARGVTYQDNA 213
QY 134 QRETEPRHMQTCQRCER-RIYEKKRQOKRYEQOREDEDEKYEERMKKEEDNKRDPQQR 192
DB 214 IRDDELPTNYTSMQKYDYVKEDKERRMQEEEOERLREYEQIOEKLQOEERERQEOE 273
QY 193 EYEDCRRCQOEPRQOHQCOLRCREOOROH 223
DB 274 RRQKRRRMEQERIRQEYD---EKERQRY 300

Search completed: March 1, 2001, 15:49:22
Job time: 367 sec

